Figure 1: Genome Threader output for INSP163 polypeptide sequence (SEQ ID NO: 30)

Genome Threader results - Energy Scores

Num	PDB Code	Norm Align Score	Raw Align Score	%IDs	%Struct Aligned	%Query Seq Aligned	Pairwise Energy	Solvation Energy	Neural Net Score	%Confidence	From Pos for Query	To Pos for Query	From Pos for Target	To Pos for Target	Alignment Length
1	1gr3A00 biopendium (align)	67.63	96	19.8	75.0	37.7	-155.89	-4.34	0.983	96	187	300	31	129	116
2	1c28A00 biopendium (align)	65.86	94	17.4	72.8	37.7	-105.30	-2.65	0.983	96	187	300	31	121	115
3	1c28C00 biopendium (align)	51.74	79	20.2	70.6	37.7	-9.01	-3.13	0.977	92	187	300	28	99	114
4	1c8BB00 biopendium (align)	45.38	73	18.2	29.7	10.9	-6.72	-2.53	0.971	88	187	219	25	57	33
5	1tsvA00 biopendium (align)	42.14	71	20.3	82.4	39.1	-94.33	1.68	0.965	84	182	299	25	146	128
6	2mtaA00 biopendium (align)	39.14	68	17.8	81.8	39.1	-89.55	1.50	0.961	81	182	299	26	146	129
7	1kxqA00 biopendium (align)	34.25	63	19.0	77.8	38.7	-43.14	3.51	0.937	69	183	299	30	141	121
8	1d2qA00 biopendium (align)	32.56	61	16.8	75.4	39.4	-12.73	4.44	0.943	72	189	301	34	134	119
9	1dq6A00 biopendium (align)	31.11	60	15.3	80.5	39.7	-7.91	-3.18	0.933	68	182	301	30	149	131

Figure 2

```
  1   tgagccgcct  cgggacggag  ccatgcggcg  ctgggcctgg  gccgcggtcg  tggtcctcct
                              m    r    r    w    a    w    a    a    v    v    v    l
      ──────────────────────────────────────────────▶
                  INSP163-CP1

 61   cgggccgcag  ctcgtgctcc  tcggggggcgt  cggggcccgg  cgggaggcac  agaggacgca
      l    g    p    q    l    v    l    l    g    g    v    g    a    r    r    e    a    q    r    t

121   gcagcctggc  cagcgcgcag  atcccccaa  cgccaccgcc  agcgcgtcct  cccgcgaggg
      q    q    p    g    q    r    a    d    p    p    n    a    t    a    s    a    s    s    r    e

181   gctgcccgag  gcccccaagc  catcccaggc  ctcaggacct  gagttctccg  acgcccacat
      g    l    p    e    a    p    k    p    s    q    a    s    g    p    e    f    s    d    a    h

241   gacatggctg  aactttgtcc  ggcggccgga  cgacggcgcc  ttaaggaagc  ggtgcggaag
      m    t    w    l    n    f    v    r    r    p    d    d    g    a    l    r    k    r    c    g

301   caggacaag  aagccgcggg  atctcttcgg  tccccagga  cctccaggtg  cagaagtgac
      s    r    d    k    k    p    r    d    l    f    g    p    p    g    p    p    g    a    e    v

361   cgcggagact  ctgcttcacg  agtttcagga  gctgctgaaa  gaggccacgg  agcgccggtt
      t    a    e    t    l    l    h    e    f    q    e    l    l    k    e    a    t    e    r    r

421   ctcagggctt  ctggaccccgc  tgctgcccca  ggggggcgggc  ctgcggctgg  tgggcgaggc
      f    s    g    l    l    d    p    l    l    p    q    g    a    g    l    r    l    v    g    e

481   ctttcactgc  cggctgcagg  gtccccgccg  ggtggacaag  cggacgctgg  tggagctgca
      a    f    h    c    r    l    q    g    p    r    r    v    d    k    r    t    l    v    e    l

541   tggtttccag  gctcctgctg  cccaaggtgc  cttcctgcga  ggctccggtc  tgagcctggc
      h    g    f    q    a    p    a    a    q    g    a    f    l    r    g    s    g    l    s    l

601   ctcgggtcgg  ttcacggccc  ccgtgtccgg  catcttccag  ttctctgcca  gtctgcacgt
      a    s    g    r    f    t    a    p    v    s    g    i    f    q    f    s    a    s    l    h

661   ggaccacagt  gagctgcagg  gcaaggcccg  gctgcgggcc  cgggacgtgg  tgtgtgttct
      v    d    h    s    e    l    q    g    k    a    r    l    r    a    r    d    v    v    c    v

721   catctgtatt  gagtccctgt  gccagcgcca  cacgtgcctg  gaggccgtct  caggcctgga
      l    i    c    i    e    s    l    c    q    r    h    t    c    l    e    a    v    s    g    l

781   gagcaacagc  agggtcttca  cgctacaggt  gcaggggctg  ctgcagctgc  aggctggaca
      e    s    n    s    r    v    f    t    l    q    v    q    g    l    l    q    l    q    a    g

841   gtacgcttct  gtgtttgtgg  acaatggctc  cggggccgtc  ctcaccatcc  aggcgggctc
      q    y    a    s    v    f    v    d    n    g    s    g    a    v    l    t    i    q    a    g

901   cagcttctcc  gggctgctcc  tgggcacgt
      s    s    f    s    g    l    l    l    g    t
           ◀──────────
              INSP163-CP2
```
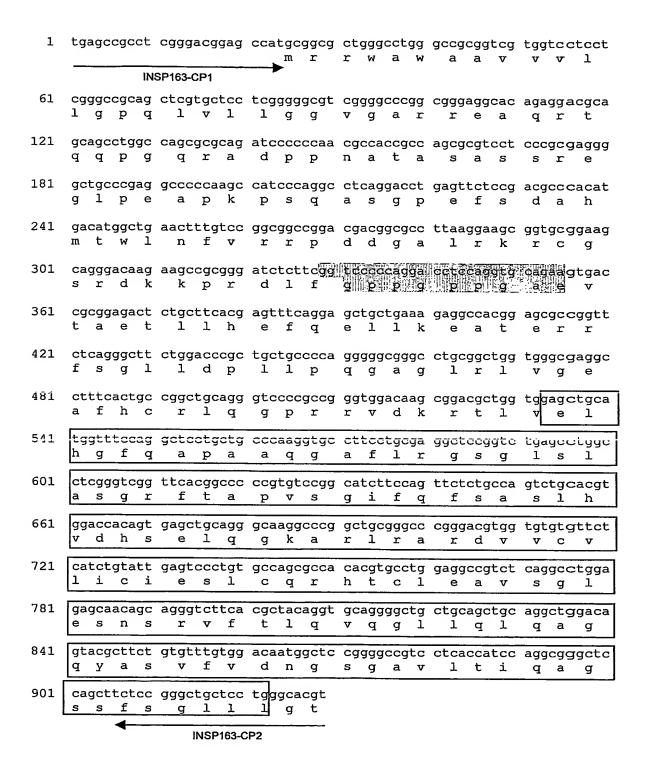
Figure 3

A) INSP163-A

```
  1    PDDGALRKRC  GSRDKKPRDL  FGPPGPPGAE  VTAETLLHEF  QELLKEATER
 51    RFSGLLDPLL  PQGAGLRLVG  EAFHCRLQGP  RRVDKRTLVE  LHGFQAPAAQ
101    GAFLRGSGLS  LASGRFTAPV  SGIFQFSASL  HVDHSELQGK  ARLRARDVVC
151    VLICIESLCQ  RHTCLEAVSG  LESNSRVFTL  QVQGLLQLQA  GQYASVFVDN
201    GSGAVLTIQA  GSSFSGLLLG  T
```

B) INSP163-B

```
  1    KRCGSRDKKP  RDLFGPPGPP  GAEVTAETLL  HEFQELLKEA  TERRFSGLLD
 51    PLLPQGAGLR  LVGEAFHCRL  QGPRRVDKRT  LVELHGFQAP  AAQGAFLRGS
101    GLSLASGRFT  APVSGIFQFS  ASLHVDHSEL  QGKARLRARD  VVCVLICIES
151    LCQRHTCLEA  VSGLESNSRV  FTLQVQGLLQ  LQAGQYASVF  VDNGSGAVLT
201    IQAGSSFSGL  LLGT
```

C) INSP163-C

```
  1    CGSRDKKPRD  LFGPPGPPGA  EVTAETLLHE  FQELLKEATE  RRFSGLLDPL
 51    LPQGAGLRLV  GEAFHCRLQG  PRRVDKRTLV  ELHGFQAPAA  QGAFLRGSGL
101    SLASGRFTAP  VSGIFQFSAS  LHVDHSELQG  KARLRARDVV  CVLICIESLC
151    QRHTCLEAVS  GLESNSRVFT  LQVQGLLQLQ  AGQYASVFVD  NGSGAVLTIQ
201    AGSSFSGLLL  GT
```

D) INSP163-D

```
  1    FSGLLDPLLP  QGAGLRLVGE  AFHCRLQGPR  RVDKRTLVEL  HGFQAPAAQG
 51    AFLRGSGLSL  ASGRFTAPVS  GIFQFSASLH  VDHSELQGKA  RLRARDVVCV
101    LICIESLCQR  HTCLEAVSGL  ESNSRVFTLQ  VQGLLQLQAG  QYASVFVDNG
151    SGAVLTIQAG  SSFSGLLLGT
```

E) INSP163-E

```
  1    VDKRTLVELH  GFQAPAAQGA  FLRGSGLSLA  SGRFTAPVSG  IFQFSASLHV
 51    DHSELQGKAR  LRARDVVCVL  ICIESLCQRH  TCLEAVSGLE  SNSRVFTLQV
101    QGLLQLQAGQ  YASVFVDNGS  GAVLTIQAGS  SFSGLLLGT
```

F) INSP163-F

```
  1    TLVELHGFQA  PAAQGAFLRG  SGLSLASGRF  TAPVSGIFQF  SASLHVDHSE
 51    LQGKARLRAR  DVVCVLICIE  SLCQRHTCLE  AVSGLESNSR  VFTLQVQGLL
101    QLQAGQYASV  FVDNGSGAVL  TIQAGSSFSG  LLLGT
```

Figure 4